

# Extensible Markup Language: An Enterprise Integration Web Delivery System

**Emmanuel Opara\***

College of Business, Prairie View A&M University, Prairie View Texas 77446  
Phone: (936) – 857- 4011, Fax: (936) – 857 - 2105, Emannual\_opara@pvamu.edu

**Mukesh Srivastava**

College of Business, Prairie View A&M University, Prairie View Texas 77446  
Phone: (936) – 857- 4011, Fax: (936) – 857 - 2105, Mukesh\_Srivastava@pvamu.edu

## ABSTRACT

*This study investigates XML technology as a standard format for data transfer on the Web. It describes several compelling reasons why XML is a particularly useful technology for representing structured systems management information. XML is a markup language used for data structure in a textual form. A specific goal of XML is to keep most of the descriptive power of SGML, while removing as much of the complexity as possible. XML is similar in concept to HTML, but whereas HTML is used to convey graphic information about a document, XML is used to represent structured data in a document. It is no wonder, global businesses are rushing to implement XML technology.*

## INTRODUCTION

Web information systems are revolutionizing commerce. For starters, the Web reduces to nearly nothing many marginal costs of doing business, such as communication and customer service. Companies such as Amazon.com and CDNow have outperformed leading national chains by exploiting the advantages of e-commerce. Web-based banks, stock traders, insurance sectors have achieved competitive advantage through the deployment of XML technologies.

According to Dodds [2000], the Web has placed in the hands of businesses, the potential to communicate with each other in real time. Many companies believe that the use of XML will enable businesses to realize many of the benefits in the area of data exchange that HTML provided for thin client, platform independent end user interface as described by Agile [2001]. Holland [2001], and Ravis and Dick [2000] noted that some early misconceptions in the industry had XML replacing HTML; however most industry leaders see them as complementary technologies. For B2B businesses to fully utilize the potentials of Web technologies, the Web will depend on widespread use of standards, because, as with the telephone, this communication medium is based on numerous layers of interoperating technology. Chang and Harkey [1998], Dick [2000], Didie [2000], Holland [2000], among others noted that such important layer is necessary because of its

visual display and user interface capabilities, that is exemplified by standards such as HTML, SGML, GIF [Graphic Interchange Format] and Live Script [previously JavaScript]. These standards allow a page to be created once, and displayed at different times by many receivers reported by Travis [2000].

Although visual and user interface standards are a necessary layer, Walsh [1998], XMI/EDI Group [1998], and Hollander [1999], found evidence that the visual and user interface standards are insufficient for representing and managing data. According to Boumphrey [2000], today's Internet is merely an access medium to text and pictures. There are no adequate standards for intelligent search, data exchange, adaptive presentation, and personalization. The Internet platform is expected to exceed the medium of merely setting an information access and display standard. It should include an information understanding standard or platform which can be explained as a standard way of representing data so that software can better search, move, display, and manipulate information currently hidden in contextual obscurity [Travis 2000].

Earlier studied by Walsh, [1998], and Hollander [1999] found that HTML provide rich facilities for data display, but does not provide any standards-based way to manage data as data. More recent studies have paid considerable attention to the need for a standard for data representation that will expand the Internet in much the same way that the HTML standard display did a few years ago. The data standard will be the medium for B2B business transactions, publication of personal preference profile, automated collaboration, and database sharing. Chen [2000] noted that payments of bills via connectivity, medical histories, pharmaceutical research data, semi-conductor part sheets, and purchase orders will be written in this format. Furthermore, this medium will open up a wide variety of new users, all based on a standard representation for moving structured data around the Web as easily as HTML pages are moved today as per Sun Microsystems [2001]. Agital [2001], Boumphrey and Tittel [2000], Dodds [2000], Dick [2000], among others found evidence that the data standard in question is eXtensible Markup Language [XML].

XML stands for eXtensible Markup Language. It is a subset of the Standardize Generalized Markup Language [SGML]. XML is designed to describe data using a Document Type Definition [DTD] or Schema. Markup Languages [such as XML or SGML] are designed to add structure and convey information about document and data. In "markup" languages, the main mechanism for supplying structural and semantic information is by decorating the document with "elements" comprising a "start tag", optionally some content, and an "end tag". Example is <tag> data</tag>. XML has been standardized by the World Wide World Consortium [W3C]

The question now is "How extensive is the use of XML, how is it being adopted and what are the set backs? Can B2B businesses use this tool as a strategy for competitive advantage"?

## **PURPOSE OF THIS STUDY**

This study attempts to answer the above questions and further examine how XML can be used as a standard format for Web information. Further, the study shows how XML will benefit B2B businesses seeking to develop integrated business solutions for the 21<sup>st</sup> century. Also, to better understand the current state of XML adoption and uncover issues that global businesses face in XML standardization.

## **RESEARCH METHOD**

Survey instruments were designed and conducted in randomly selected industries in selected countries. The survey instrument was designed using questionnaires that covered the following variables: XML development, realm of XML standardization, adoption decision criteria used, organizational policies on IT, integration of a Straight-through-Processing [STP] technology, acceptance and belief in Reduced Settlement Cycle [T+1] process for a standard data-exchange format that is based on XML, management commitment to the process, involvement in the development of company information technology [IT] strategies, role of Internet technologies such as SGML, HTML, Electronic Data Interchange [EDI], Simple Object Access Protocol [SOAP] and XML on company's Web data transfer, adoption of XML technology enterprise-wide.

Sample population is composed of eleven companies in the United States, Canada, Australia, Germany, Great Britain, Switzerland, Korea, Norway, Netherlands, Sweden and Spain. The targeted companies were selected from eight industry sectors comprising of technology, oil and gas, pharmaceutical, communication, banking and finance, investments, automotive, consumer food. Survey questionnaires were translated into languages by translating agencies in the countries surveyed. The survey targeted information-technology executives, business executives and consultants in areas ranging from technology, investment banking and securities, to funds and asset management, communication, and medical. Company sizes ranged from annual revenues of less than \$8 million to fortune 500 companies. Telephone interviews were also performed with information technology executives and system managers responsible for IT decisions in their areas. The interviews were carried out by professional assisted call centers. In total, 2400 companies were surveyed or interviewed. These countries and companies were selected in cooperation of Shell Deep Water Services – Niger Delta Team [a subsidiary of Shell Oil Corporation], who rendered help in survey translation, distribution and collection from participants in oversea countries. For each country and respective industry set, an average of 40 companies was interviewed. They were randomly selected from a local industry database. This process ensured that, for each country, the total sample size exceeds N=200, and for each industry, the sample size exceed N=300.

## **RESULTS**

According to the survey findings, the adoption of XML technology as the emerging standard for exchanging and presentation of data on the Internet has become an important business initiative in the 21<sup>st</sup> century. The result showed that XML has become a

strategic enterprise-wide business initiative driven by information technology and business executives. The survey also found that while the majority of XML adoption initiatives are being sponsored by IT groups and departments, 33% of those initiatives were driven by top executives' management. Furthermore, the findings found that more than 82% of those surveyed considered XML to be critical to the company's enterprise-wide long term IT strategy. On the question of XML implementation, 45% of the respondents noted that they have already implemented an XML strategy, while 38% are planning to get into the implementation phase by next year. However, 17% of the respondents were not sure and are waiting to see how those who have already implemented an XML strategy have performed on their e-commerce ventures.

On survey question on the importance of XML to senior management, respondents noted that XML is important due to the financial-industry initiatives and regulations affecting the industry. They noted that these include the "Straight-Through-Processing" [STP] and "Reduced Settlement Cycles" [T+1], which has increased the need for a standard data-exchange format that is based on XML. STP technology framework seeks to provide efficiencies through seamless data flow within the enterprise as well as across the market without manual intervention. T+1 technology involve shorter settlement business cycles with regard to financial and IT operations. This technology provides accurate and timely back-office information that is needed by financial and IT clients to help limit the number of disputes in a settlement process. These technologies will fit into XML environment because of its system integration capabilities. Respondents further confirmed that industry requirements, such as STP and T+1 technology were among the most important reasons for their companies' adaptation of XML. The results further revealed that senior executives played active roles in the development and deployment of their respective companies' information technology strategies, since technology is increasingly recognized as being "critical" to the success of an organization. According to Boumphrey [2000], Internet technologies are very crucial in the overall business strategies as companies try to meet and exceed the expectations of its customers.

The survey also confirmed that technologies, like XML, are seen as important for achieving greater efficiencies and improving companies' bottom lines. Focusing on this strategy will be important for enabling the next phase of XML development that will address specific vertical markets.

On the question of technology developments and initiatives that impact the capital markets, the survey result showed some interesting developments in the area of XML standardization. Since 82 % of those surveyed were committed to XML as a business initiative, it is obvious that XML strategies are important as a competitive technological tool in global e-business. These strategies are expected to continue to grow in complexity as e-transaction processing in the banking-financial services industry leans on IT to ensure compliance and maintain an above average returns on investments by Dodds [2000].

On the survey question of the importance XML for Enterprise Systems Integration, the result showed that global companies are recognizing the importance of XML as a key

technology for enterprise systems-integration strategy. Among the group that responded to this question [the importance XML for Enterprise Systems Integration], more than half believed that XML is an important enabling technology for the integration of enterprise IT systems. Also, that the sharing of business processes standards and rules present more opportunities to leverage XML technologies.

On the survey question about where XML is currently being deployed, majority of the firms responded stating that deployment of XML can be found across their entire enterprise. The survey noted that the strategy that began in the back office for applications, such as routing FIX messages, has now accelerated to include front-office activities, such as trade-order management in the supply chain process. This is as a result of XML.

## **MARKUP LANGUAGES**

Travis [2000], Sterling Commerce [2001], Tibco [2002] among others defined a markup language as a special set of indicators, called tags that indicate how information should be interpreted. These are codes that specify how text, graphics and other elements are displayed on the Web and indicate where links lead. From a lay-man's point of view, it is similar to using a highlighter to mark an important subject matter in a text book. For example, the new office XP word processors use them to specify formatting and layout. In connectivity functions, programs use markups to express the meaning of data across networks. Enterprise integrated database systems, which must associate meaning and relationships with data they serve, and multimedia interactive processing programs, which also express meta-data about images and sound, are examples of markups. The following sections describe the history of XML, and position the technology with respect to other markup languages and applications such as SGML and HTML.

### ***Why XML?***

The strategic importance of XML is shown in its ability to be "self-describing." For example, every piece of XML data contains information that describes exactly what it does as shown by Arbortext [2001]. XML is known for the simplification of communications between non-homogenous platforms which allows the vision of the virtual supply chain and customer relationship management to become a reality.

Travis [2000], Holland [2000], Dick [2000] among others noted that XML is not new, but started in 1996 and has been a World Wide Web Consortium [W3C] standard since February 1998. It has been widely used for large documentation by Morgenthal and La Forge [2001]. Before the inception of XML, was the Standard General Markup Language [SGML], developed in the early 80s. SGML is an International Standard Organization [ISO] standard. It was widely used for large document projects. Hypertext Markup Language [HTML] evolved in the 90s and is widely popular for display data on the Web. According to Dodds [2000], the designers of XML simply integrated the best parts of SGML and HTML to develop XML.

XML was developed because of the need to represent and manage data on the Web. HTML, GIF and Jscript<sup>TM</sup> were the standard for visual display and user interface on the Web. These standards allow a page to be created once and be displayed at different times by the receivers. Representation and management of data between different platforms was problematic.

HTML, for example, does not support a common way of representing data so that software can search, move, display and manipulate data. Secondly, XML and XML extensions have now become the standard for Web-based data transactions such as managing e-commerce business processes that connect the B2B businesses with its clients.

Recent studies have shown that B2B businesses have begun developing XML-based voice files to deliver information through wireless phones to its clients as reported by Halland [2002], TIBCO [2002], Didie [2000], Arciniegas [2000]. For example, Detroit-based General Motors Corporation has claimed that they will be providing OnStar XML-based wireless voice files to deliver weather forecasts, news sports scores and stock updates to wireless phones built into 30 models of their premier 2002 cars according to GMC Annual Report [2001].

According to Chen [2000], if a business strategy is to deliver Web content to devices, then they need to be using XML and XSL [Extensible Stylesheet Language] because 86% to 94 % of firms are committed to using this strategy to attain competitive advantage.

Agital define XML as a subset of SGML optimized for delivery for delivery of data over the Web as illustrated by W3C. Dodds [2000] noted that XML improved over some of the deficiencies found in SGML. The deficiencies include legacy requirements and features that make generating SGML-based documents easier, but complicating on its use on the Web.

XML offers interoperability using a flexible, open standard giving B2B client's new ways to access and deliver data. It ensures that structured data will be uniform and independent of applications or vendors. In industrial automation, XML is the driving force in data exchange, manipulation, and transmission. XML, in collaboration with HTML and Simple Object Access Protocol [SOAP], will provide the foundation for Web-Based Human Machine interface (Travis 2000).

### ***XML and HTML; Is there a difference***

According to Dick [2002], HTML is about user interface while XML is about data. Dynamic HTML describes display and user interaction. Didie [2000] noted that XML describes information. This can be supported in the argument, that XML can add information to an HTML document and that HTML can display information expressed in XML format. Table 1 explains some of the differences between the XML and HTML.

Table 1 - XML versus HTML: Is there a difference

<b>XML</b>	<b>HTML</b>
Best for data Structure	Used for display purposes
Presentation independent	No knowledge of data
Open Language	Closed Language; Standard
Case sensitive {due to the integration of Unicode}	Is not case sensitive
Tags without content and those which are empty elements must use “/>” notation	Empty tag like   requires nothing
White space, including line breaks, is significant within content	Special white space is ignored

Source: XML/EDI Group (2001)

### ***Compatibility of HTML with XML***

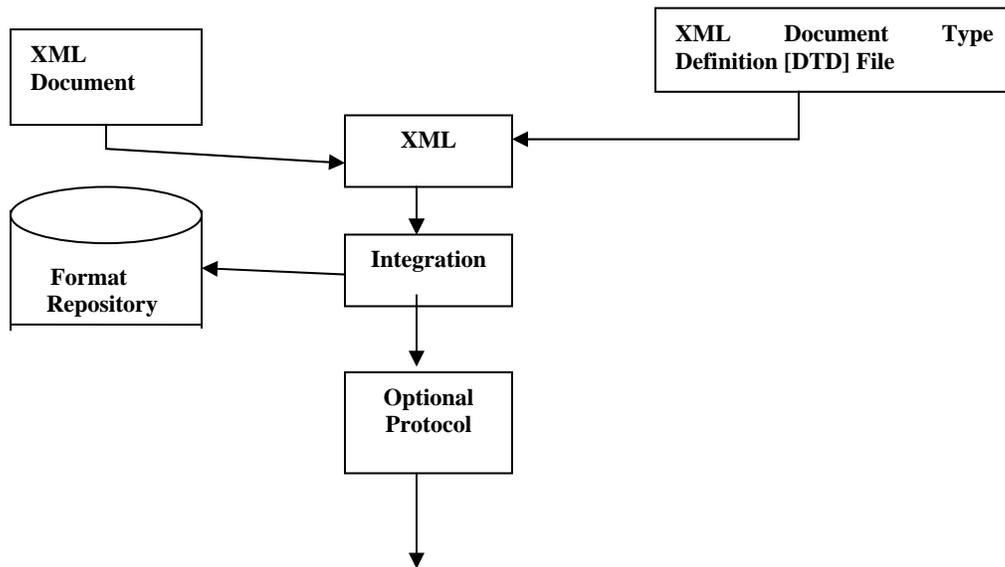
Complaint about HTML is that it is not extensible. The set of tags is fixed and fairly display-centered, which makes it difficult to add information such as revision histories or to mark-up displayed text. Dodds [2000] noted that it is not easy to add semantic information to HTML pages. HTML is not easily extended for data representation, due to its nature as a display language and partly because it was not for open extensibility. To resolve this problem, Microsoft is coordinating with the W3C to define a format for putting XML data inside HTML pages. By extending HTML to allow arbitrary XML data elements, a wide range of applications can use HTML as the primary document or display format, and also use XML embedded within these documents to hold application-specific data.

### ***Modeling for the Future***

XML is the technology for Web interactive processes and is growing in importance in the industrial automation world from embedded systems to human machine interface, to distributed systems and database management. It is used as a data definition and manipulation language. Since data is being exchanged in the XML format, it can be easily integrated from different platforms. An example [see diagram 1], using Microsoft's

BizTalk specification that consist of a framework description, a repository for integrated-industry schemas, and a process for posting and validating schemas. The diagram shows how XML communicate with other software packages.

**Diagram 1: BizTalk Specification on XML**

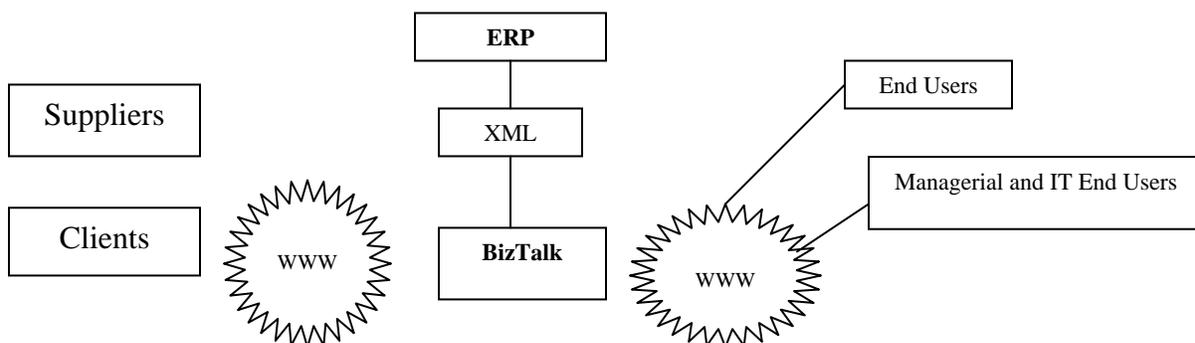


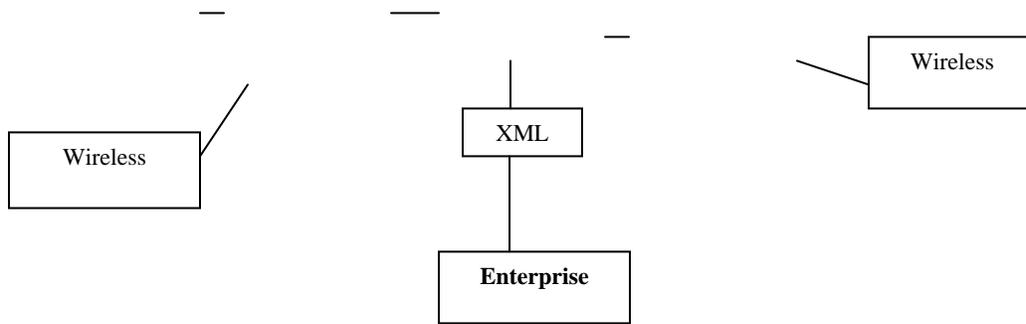
Source: Travis (2000)

The BizTalk server provides B2B document interchange and provides a server for routing information between clients, suppliers and customers based on the specific value (Travis 2000). Enterprise Resource Planning [ERP] leading technology companies such as SAP, BAAN, Peoplesoft, JD Edward, are developing Application Integrated Component [AICs] as shown in diagram 2. These AICs provide the capability for software to integrate with Microsoft’s BizTalk. For example, an alarm occurs that generate an XML-based BizTalk’s representation of a report.

This is sent to a BizTalk server, and the BizTalk Server then shares the information with other application or the Web [see diagram 2]. The BizTalk Server is a scalable product that will facilitate the interchange of BizTalk Framework-encoded information

**Diagram 2: Application Integration Component [AICs]**





Source: Travis (2000)

## IMPLICATION TO MANAGERS AND PRACTITIONERS

XML is particularly useful in certain circumstances and scenarios. As a universal standard for the expression of data, it provides a data standard that can encode the content, semantics, and schemata for gamut of cases, from simple to complex. The flexibility of a single data representation format allows any software to determine the semantics of a data element, without previous knowledge of the underlying meaning of the data by Goldfarb and Prescod [2000]. The effect of that is that information can then be reused for new purposes and in novel contexts.

For software developers building Web applications and line-of- business Internet software, XML provide a powerful, flexible format for expressing data. This format could be a wire format for sending data between client and server, a transfer format for sharing data between applications, or a persistent storage format on disk.

XML offers a mechanism for adding meta-data or meta-content to HTML to a Web site. For end-users, it promises to provide a much richer set of Web applications for browsing, communication, and collaboration by Boumphrey [2000].

### ***Implication for Researchers***

The technology development cycle for XML is now the “excitement phase”. What does the future hold? Addition research will be done to answer that question. However, the next few years will show the impact of XML on e-commerce as new information retrieval capabilities will be possible from an XML enabled Web base infrastructure. Also, with the increased capabilities of businesses to store and process data, the role of the browser will change from one presentation to that of being an application tool. This change will result from the combination of XML based data and HTML embedded scripting or Java to customize the presentation to the user’s needs. Further research will confirm all the mentioned possibilities.

## CONCLUSION

The Study has argued that XML offers interoperability using a flexible, open standard and redefined breakthroughs to new ways to access and deliver data. We have proposed that XML will benefit B2B businesses seeking to develop integrated business solutions in the 21<sup>st</sup> century. XML came about because of the need to represent and manage data on the Web. Holland [2000], Chen [2000], Dick [2000] among others found evidence that XML is growing in importance in the industrial automation world from embedded systems to human machine interface, to distributed systems and database management. Agital [2001] noted that XML has become the de-facto standard for data communication between different applications, systems and through the Internet. Extensible Markup Language [XML] offers interoperability using a flexible, open standard giving client's new ways to access and deliver data. The study further concludes by citing that the continuous growth in the use of XML will improve Web-browsing applications for viewing, filtering, and manipulating information on the Internet. In industrial automation, XML has been the driving force in data exchange, manipulation, and transmission. XML, in collaboration with HTML and SOAP, will continue to provide the foundation for Web-Based Human Machine interface.

From a strategic point of view, the real advantage of XML will be when the industrial automation world, standardizes on the vocabularies that allows problem reports, suppliers request, part numbers, customer numbers, document types, shopping characteristics, and all the terminology that allows B2B businesses, to work together and communicate from the factory floor to the top floor using the same language. XML implementation will eventually lead to cost savings, interoperability and new opportunities for businesses. As collaboration on the Web spreads to more businesses, customer services will eventually migrate from phone lines and storefronts to Web sites [Oracle 2001]. The majority of these Intranet and Extranet applications will involve manipulation or transfer of data and database records, such as purchase orders, invoices, customer information, appointments, [Kleinman 2001]. As noted by IT executives at Sun Microsystems, the XML technology will continue to enhance the richness of end-user capabilities on the Web because it enables such a wide array of business applications to be implemented on the Internet.

## BIBLIOGRAPHY

A., Fabio Arciniegas, "What is XLink?"

Agital. (2001, August). *XML e-business solution available for free 30-day test.*

<http://www.agital.net/company/press>. (retrieved November 25, 2001), 2001

Arbortext Epic E-Content Engine and Empolis Sigmalink integrate for XML content management and publishing solution. (2000). *Arbortext News and Events*. [http://www.arbortext.com/News\\_and\\_Events/recent\\_news\\_pages](http://www.arbortext.com/News_and_Events/recent_news_pages). (retrieved November 25, 2001), 2000.

Arbortext. (2001). XML for managers. Arbortext Think Tank.

[http://www.arbortext.com/Think\\_Tank/XML\\_for\\_Managers/xml\\_for\\_managers.html](http://www.arbortext.com/Think_Tank/XML_for_Managers/xml_for_managers.html). (retrieved September 10, 2001), 2001.

- Boumpfrey, F., & Tittel, E. *XML for Dummies*. Foster City, CA: IDG Books Worldwide, 2000.
- Boumpfrey, F.; Dizenzo, O.; and et. Al., *XML Applications*. USA: Wrox Press, 1998.
- Boumpfrey, F.; Greer, C.; and et. Al., *Beginning XHTML*. USA: Wrox Press, 2000.
- Ceponkus, A., and Hoodbhoy, F., *Applied XML: A Toolkit for Programmers*. United States: John Wiley & Sons, 1999.
- Chang, D., and Harkey, D., *Client/Server Data Access with Java and XML*. United States: John Wiley & Sons, 1998.
- Chen, A., "Getting to XML: The new B2B lingua Franca will need to live with EDI. Here's how to mix the two." eWeek <http://www.zdnet.com/eweek/stories/general/0,11011,2610043,00.html> (7 August 2000), 2000.
- Dick, K., *XML A Manager's Guide*. Reading, MA: Addison-Wesley, 2000.
- Didie, M. "Getting into i-Mode." <http://www.xml.com/pub/2000/09/20/wireless/imode.HTML> (retrieved 20 September 2000), 2000.
- Dodds, L., "Gentrifying the Web: Taking a look at XHTML, the XML-Deviant finds that although the W3C HTML Activity is moving forward, the rest of the web is still lagging behind." <http://www.xml.com/pub/2000/09/13/xhtml/index.html> (retrieved 13 September, 2000), 2000.
- Dodds, L., "Schema Round-up." <http://www.xml.com/pub/200/09/06/deviant.html> (retrieved 6 September, 2000), 2000.
- General Motors Annual Report. 2001.
- Goldfarb, C.F., & Prescod, P. *The XML Handbook*. Upper Saddle River, NJ: Prentice Hall, 2000.
- Holland, R., "More and more, XML diving into the 'stream.'" eWeek <http://www.zdnet.com/eweek/stories/general/0,11011,2608541,00.html> (27 July 2000), 2000.
- Holland, R., "XML bridges e-apps." eWeek <http://www.zdnet.com/eweek/stories/general/0,11011,2585447,00.html> (7 August 2000), 2000.
- Hollander, J., "Netscape Sees XML in Its E-Commerce Future." E-Commerce Times <http://www.ecommercetimes.com/news/articles/990309-2.shtml> <http://www.xml.com/pub/2000/09/xlink/index.html> 18 September 2000, 1999.
- Morgenthal, J. P. and La Forge, B. *Enterprise Application Integration with XML and Java*. Upper Saddle River, NJ: Prentice Hall, 2001.
- Oracle. *Using XML in Oracle database applications*. <http://technet.oracle.com/tech/xml/info/htdocs/otnwp>. (retrieved November 25, 2001), 2001.
- Sterling Commerce. *XML and e-Business*. <http://www.sterlingcommerce.com/solutions/products/ebi/wp/xmlbiz%5Fwp.html>. (retrieved September 25, 2001), 2001.

- Sun Microsystems. *Sun Microsystems leads industry effort to simplify XML development*. <http://www.sun.com/smi/Press/sunflash/1999-12/sunflash.991207.2.html>. (retrieved November 25, 2001), 1998.
- TIBCO. *TIBCO Extensibility*. TIBCO Software, Inc. <http://www.tibco.com/solutions/products/extensibility>. (retrieved September 2, 2002), 2002.
- Travis, B., *XML and SOAP Programming for BizTalk Servers*, Redmond, Washington, Microsoft Press, 2000.
- Walsh, N. "A Technical Introduction to XML." <http://www.xml.com/pub/98/10/guide0.html> (retrieved 3 October 1998), 1998.
- XML/EDI Group. *Guidelines for using XML for electronic data interchange*. <http://www.xmledi-group.org/xmledigroup/guide.htm>. (retrieved September 24, 2001), 1998.